

AN AUTOMATIC METHOD OF IDENTIFYING IMAGE SUBREGIONS FOR REUSE DURING DATASTREAM TRANSMISSION

BACKGROUND OF THE INVENTION

1. Technical Field:

The present invention relates in general to data transmission and in particular to reducing datastream transmission bandwidth. Still more particularly, the present invention relates to reducing the bandwidth consumed in transmission of a datastream by encoding image data within the datastream through a compression algorithm.

2. Description of the Related Art:

The revolution in data processing system speed has brought about a corresponding expansion of the types of tasks performed by data processing systems. In the field of image processing, data processing systems now receive, process, and produce all manner of images and image-embedded documents. Many, if not most, of these tasks involve the transmission of images between a data processing system and a peripheral device or another data processing system. Examples of such tasks include the printing of a paper document, the acquisition of a digital image from a digital camera, and the viewing of a web page.

The capacity of interlinks between data processing systems, frequently called available bandwidth, has improved at a substantial pace, and the capacity of interlinks between data processing systems and their peripherals has also improved, although at a frustratingly slow pace. Because of the long process of standardization in peripheral interfaces and protocols for machine interaction, revolutionary improvements in available bandwidth have arrived only very slowly. The traffic across the available bandwidth in many applications, however, has increased at a much faster pace.

Digital cameras provide an excellent example of this phenomenon. The last few years have witnessed a tremendous improvement in the pixel resolution of the images captured by digital cameras, but this revolution in resolution has exponentially increased the size of the image files that must be transported between the camera peripheral and the data processing system that processes the image files. The stagnation in capacity of the available interlinks has created tremendous frustration among users as they wait for images to upload from the digital camera to the data processing system.

Other examples include the increasing reliance on images in ever more sophisticated desktop publishing applications, and the ubiquitous images that have turned the worldwide web from a text-based interface to a complicated multimedia experience. In the case of the worldwide web, it was once predicted that the images embedded in web pages would eventually bring the internet to a standstill. Though the prophets of doom predicted an outcome far more bleak than reality, the frustration of waiting for large images embedded in web pages to download is a disturbingly common experience, especially for users of dial-up modems, which communicate over conventional telephone lines.

With no immediate hope of expanding the bandwidth available for the transmission of image data between data processing systems and between data processing systems and their peripherals, and with the volume of image data transmitted increasing almost daily due to improvements in image acquisition, image processing, and data storage, what is needed is a method of reducing the bandwidth consumed in transmission of a datastream by encoding image data within the datastream through a compression algorithm. Such a method would reduce perceived delay in transmissions and correspondingly reduce the frustration of users who wait impatiently as image content is delivered to their data processing systems, their printers, or their web terminals.

SUMMARY OF THE INVENTION

It is therefore one object of the present invention to improve the efficiency of data transmission.

It is another object of the present invention to reduce datastream transmission bandwidth.

It is yet another object of the present invention to reduce the bandwidth consumed in transmission of a datastream by encoding image data within the datastream through a compression algorithm.

The foregoing objects are achieved as is now described. A method, system, and computer program product are provided, wherein the bandwidth necessary to transmit an item of image data is reduced. When items of image data are identified in a datastream, they are extracted from the datastream. The image data item is then divided into a series of subregions of variable size. Where efficiency dictates that the operation is appropriate, the subregions are replaced in the image data item with a unique identifier to produce a reduced image. The reduced image is then packaged into a new data structure containing a header, the reduced image, and a decoding table that will allow the replacement of the identifiers with the extracted subregions. Where subregions are repeated, as they frequently are in images of large size, this arrangement will allow for the compression of the image by the elimination of redundant data that merely represents a repeated subregion. When the image reaches its destination, the image is decoded to reproduce the original image.

BRIEF DESCRIPTION OF THE DRAWINGS

The above as well as additional objects, features, and advantages of the present invention will become apparent in the following detailed written description.

The novel features believed characteristic of the invention are set forth in the appended claims. The invention itself however, as well as a preferred mode of use, and further objects and advantages thereof, will best be understood by reference to the following detailed description of an illustrative embodiment when read in conjunction with the accompanying drawings, wherein:

Figure 1 depicts a representative example of a networked environment, containing multiple data processing systems and peripherals, in which a preferred embodiment of the present invention may be implemented;

Figure 2 is a schematic representation of the decoded content of an image data structure in accordance with a preferred embodiment of the present invention;

Figure 3 depicts a simplified representation of a packaged image data structure in accordance with a preferred embodiment of the present invention;

Figure 4 is a high-level flowchart for a process of encoding an image data structure to reduce the bandwidth consumed in transmission of a datastream by employing a process of encoding image data within the datastream through a compression algorithm in accordance with a preferred embodiment of the present invention;

Figure 5 depicts a high-level flowchart for a process of decoding a packaged image data structure to reduce the bandwidth consumed in transmission of a datastream by employing a process of encoding image data within the datastream through a compression algorithm in accordance with a preferred embodiment of the present.

5

Figure 6 is a module diagram representing message flow between the various software modules in a process of encoding image data within a datastream through a compression algorithm in accordance with the preferred embodiment of the present invention.

10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

With reference now to the figures, and in particular with reference to **Figure 1**, a representative example of a networked environment, containing multiple data processing systems and peripherals, in which a preferred embodiment of the present invention may be implemented, is illustrated. The environment contains a first data processing system **100** and a second data processing system **102**, which are linked across a network **104**. First data processing system **100** is attached to a first display device **106** across a direct inter-system interface **108**, such as a cable. An additional display device **110** is also attached to network **104**. First display device **106** and additional display device **110** discussed in this exemplary embodiment are printers interacting with both a first data processing system **100** and a second data processing system **102**, but the choice of printers and data processing systems is merely exemplary and is not meant to limit the scope of the invention. The present invention applies to any devices that process image data items. These may also include, without limitation, printing systems, data processing systems, personal digital assistants, image reproduction systems, and image display systems of all types. As is shown in the preferred embodiment, linkage across a network **104** is optional and the invention can be practiced and used in an environment where devices connect directly across a direct inter-system interface **108**.

First data processing system **100** and second data processing system **102** are represented in a highly simplified manner, and many possible components, which are not critical to understanding the present invention and are familiar to those skilled in the art, are omitted for the sake of clarity. First data processing system **100** contains a processor **112**, a RAM unit **114**, a fixed disk storage unit **116**, a user I/O controller **118**, a network interface **120**, and a print controller **122**.

RAM unit 114 serves as a short term storage location for stored data 115 and instructions as processor 112 operates on stored data 115 and instructions. In the preferred embodiment of the present invention, RAM unit 114 is shown as containing multiple programs and a data structure of stored data 115. The programs include an operating system 124 and a user application 126. Though only those programs are shown in RAM unit 114 of the preferred embodiment, many additional programs, which are omitted for the sake of simplicity and clarity, may also run on first data processing system 100 without departing from the scope or spirit of the present invention. Operating system 124 will typically control program execution, resource allocation, input/output operations, and other functions of first data processing system 100. It exists as a series of modules, only two of which are shown for the sake of simplicity. I/O module 128 controls the instructions sent to user I/O controller 118, network interface 120, and print controller 122. Encoding system 130 controls and performs the image processing functions that form part of the method of reducing datastream transmission bandwidth of the preferred embodiment. The other program shown as stored in RAM unit 114, user application 126, provides a variety of functions that will vary from data processing system to data processing system without impact on the preferred embodiment, though they will often include applications that will generate transmission datastreams, such as print jobs, on which the preferred embodiment will operate.

Other components of first data processing system 100, whose functions have not yet been explained, will typically include a fixed-disk storage unit 116, a user I/O controller 118, a network interface 120 and a print controller 122. Fixed-disk storage unit 116 serves as a long term storage location for data and instructions. User I/O controller 118 provides an interface for most peripheral equipment while network interface 120 provides physical connectivity to network 104 that allows logical connectivity to second data processing system 102 and additional display device 110.

Print controller **122** provides physical and logical connectivity to first display device **106** across a direct inter-system interface **108**.

5 First data processing system connects to a first display device **106** across a direct inter-system interface **108**, such as a cable. First display device **106** of the preferred embodiment is a printer, but first display device **106** is represented in a highly simplified manner, and many possible components, which are not critical to understanding the present invention and are familiar to those skilled in the art, are omitted for the sake of clarity. The components shown as part of first display device **106** include an I/O controller **132**, a RAM unit **134**, and a printing controller **136**. I/O controller **132** provides physical and logical connectivity to first data processing system **100** across a direct inter-system interface **108**.

10 RAM unit **134** serves as a short term storage location for stored data **138** and instructions as printing controller **136** operates on stored data **138** and instructions in directing the operation of first display device **106**. In the preferred embodiment of the present invention, RAM unit **134** is shown as containing multiple programs and a data structure of stored data **138**. The programs include an operating system **140** and printing applications **142**. Though only those programs are shown in RAM unit **134** of the preferred embodiment, many additional programs, which are omitted for the sake of simplicity and clarity, may also run on first display device **106** without departing from the scope or spirit of the present invention. Operating system **140** will typically control program execution, resource allocation, input/output operations, and other functions of first display device **106**. It exists as a series of modules, only two of which are shown for the sake of simplicity. I/O module **144** controls the instructions sent to I/O controller **132** and printing controller **136**. Decoding system **146** controls and performs the image processing functions that form part of the method of reducing datastream transmission

bandwidth of the preferred embodiment. The other program shown as stored in the RAM unit **134**, printing applications **142**, provides a variety of functions that will vary from one display device to another without impact on the preferred embodiment, though they will often include applications that will receive transmission datastreams, such as print jobs, on which the preferred embodiment will operate.

Second data processing system **102** is attached to additional display device **110** and to first data processing system **100** across network **104**. Second data processing system **102** is represented in a highly simplified manner, and many possible components, which are not critical to understanding the present invention and are familiar to those skilled in the art, are omitted for the sake of clarity. Second data processing system **102** contains a processor **148**, a RAM unit **150**, a fixed disk storage unit **153**, a user I/O controller **154**, and a network interface **156**.

RAM unit **150** serves as a short term storage location for stored data **152** and instructions as processor **148** operates on stored data **152** and instructions. In the preferred embodiment of the present invention, RAM unit **150** is shown as containing multiple programs and a data structure of stored data **152**. The programs include an operating system **158** and a user application **160**. Though only those programs are shown in RAM unit **150** of the preferred embodiment, many additional programs which are omitted for the sake of simplicity and clarity may also run on second data processing system **102** without departing from the scope or spirit of the present invention. Operating system **158** will typically control program execution, resource allocation, input/output operations, and other functions of second data processing system **102**. It exists as a series of modules, only two of which are shown for the sake of simplicity. I/O module **162** controls the instructions sent to user I/O controller **154** and network interface **156**. Decoding system **164** controls and performs the image processing functions that form

part of the method of reducing datastream transmission bandwidth of the preferred embodiment. The other program shown as stored in RAM unit **150**, user application **160**, provides a variety of functions that will vary from data processing system to data processing system without impact on the preferred embodiment, though they will often include applications that will receive transmission datastreams, such as print jobs, on which the preferred embodiment will operate.

Other components of second data processing system **102**, whose functions have not yet been explained, will typically include a fixed-disk storage unit **153**, a user I/O controller **154** and a network interface **156**. Fixed-disk storage unit **153** serves as a long term storage location for data and instructions. User I/O controller **154** provides an interface for most peripheral equipment while network interface **156** provides physical connectivity to network **104** that allows logical connectivity to first data processing system **100**.

Second display device **110** connects to first data processing system **100** and second data processing system **102** across network **104**. Second display device **110** of the preferred embodiment is a printer, but second display device **110** is represented in a highly simplified manner, and many possible components, which are not critical to understanding the present invention and are familiar to those skilled in the art, are omitted for the sake of clarity. The components shown include a network interface **166**, a RAM unit **168**, and a printing controller **170**.

The RAM unit **168** serves as a short term storage location for stored data **172** and instructions as printing controller **170** operates on stored data **172** and instructions in directing the operation of second display device **110**. In the preferred embodiment of the present invention, RAM unit **168** is shown as containing multiple programs and a data

structure of stored data 172. The programs include an operating system 174 and printing applications 176. Though only those programs are shown in RAM unit 168 of the preferred embodiment, many additional programs, which are omitted for the sake of simplicity and clarity, may also run on second display device 110 without departing from the scope or spirit of the present invention. Operating system 174 will typically control program execution, resource allocation, input/output operations, and other functions of second display device 110. It exists as a series of modules, only two of which are shown for the sake of simplicity. I/O module 178 controls the instructions sent to network interface 166 and printing controller 170. Decoding system 180 controls and performs the image processing functions that form part of the method of reducing datastream transmission bandwidth of the preferred embodiment. The other program shown as stored in RAM unit 168, printing applications 176, provides a variety of functions that will vary from one display device to another without impact on the preferred embodiment, though they will often include applications that will receive transmission datastreams, such as print jobs, on which the preferred embodiment will operate.

With reference to **Figure 2**, a schematic representation of the decoded content of an image data structure in accordance with a preferred embodiment of the present invention is depicted. An unprocessed image 200, represents the decoded content of an image data structure as it would be displayed to a user of a display system. It contains a first unique text area 202, a second unique text area 204, and a third unique text area 206. It also contains a first repeated image component 208, a second repeated image component 210, and a third repeated image component 212. An image data structure as processed by the preferred embodiment of the present invention 213, shown in the same figure, represents the same decoded content of an image data structure as it would be processed by the preferred embodiment of the present invention. It contains a first unique text area 214, a second unique text area 216, and a third unique text area 218.

Because first unique text area **214**, second unique text area **216**, and third unique text area **218** are distinct from one another, they do not present an opportunity for compression under the preferred embodiment of the present invention. First repeated image component **208**, second repeated image component **210**, and third repeated image component **212**, however, present an opportunity for compression by the preferred embodiment of the present invention. The preferred embodiment of the present invention will process them as a first repeated subregion **220**, a second repeated subregion **222**, and a third repeated subregion **224**.

In the preferred embodiment of the present invention, rather than transmitting the repeated subregion three times, the transmitting system will identify the repeated subregion, transmit it once, and then substitute an identifier, indicating the proper insertion points for the repeated subregion, into the image data structure that is transmitted. In the process of, or immediately prior to, transmitting a datastream across network **104**, encoding system **130** of operating system **124** of first data processing system **100** will examine the datastream to determine if any image data structures are present. Responsive to the determination that an image data structure is present, the encoding system will extract the image data structure from the datastream. It will then divide the image data structure into subregions such as first unique text area **202**, second unique text area **204**, third unique text area **206**, first repeated image component **208**, second repeated image component **210**, and third repeated image component **212**. Seeing that first repeated image component **208**, second repeated image component **210**, and third repeated image component **212** are redundant in the datastream, encoding system **130** will then store one of these image components as a reference and assign to the selected image component an identifier. In the image data structure itself, first repeated image component **208**, second repeated image component **210**, and third repeated image component **212** will be replaced with the identifier. The encoding system will then

package the image data structure, now containing identifiers rather than the repeated image components into a data structure that is suitable for transmission.

With reference to **Figure 3**, a simplified representation of a packaged image data structure in accordance with a preferred embodiment of the present invention is illustrated. This diagram shows a packaged image data structure **300**, the result of the process described above, ready for transmission from a transmitting device, such as first data processing system **100**, to a receiving device, such as second data processing system **102**, first display device **106** or additional display device **110**. Packaged image data structure **300** will typically contain a header **302**, which will indicate to the receiving device that the data structure received is a packaged image data structure **300** in accordance with the preferred embodiment of the present invention. The header may also contain information that will be useful to the proper decoding of the packaged image data structure, such as the type and version of encoding system **130** that was used to create packaged image data structure **300**.

The packaged image data structure **300** will also contain a decoding table **304**, which will comprise a series of references and identifiers. In packaged image data structure **300** depicted in **Figure 3**, a first reference **306**, corresponding to a first identifier **308**, a second reference **310**, corresponding to a second identifier **312**, and a third reference **314**, corresponding to a third identifier **316**, are shown by way of example. Clearly, the encoding of actual images may include a potentially infinite number of references and identifiers, corresponding to a potentially infinite number of image subregions. The number of subregions that is appropriate for a particular image may be determined by the encoding system through an analysis of the most effective size of a subregion on the basis of any desired system behavior, including transmission efficiency and encoding efficiency.

The packaged image data structure 300 will also contain a reduced image 318. The reduced image will contain the raw data necessary to decode the packaged image, including a series of unencoded image portions and identifiers that will indicate places into which the references should be substituted by decoding system 164 of the receiving device. Reduced image 318 of packaged image data structure 300 shown in **Figure 3** contains a first unencoded image item 320, a first identifier 308, a second unencoded image item 322, a second identifier 312, a third unencoded image item 324, a third identifier 316, and a fourth unencoded image item 326. Clearly, the encoding of actual images may include a potentially infinite number of identifiers and unencoded image items, corresponding to a potentially infinite number of image subregions. The number of subregions that is appropriate for a particular image may be determined by the encoding system through an analysis of the most effective size of a subregion on the basis of any desired system behavior, including transmission efficiency and encoding efficiency.

When packaged image data structure 300 is decoded by the receiving system, reduced image 318 is read, and the appropriate reference is substituted for each identifier. This allows the re-creation of the original image data structure.

With reference to **Figure 4**, a high-level flowchart for a process of encoding an image data structure to reduce the bandwidth consumed in transmission of a datastream by employing a process of encoding image data within the datastream through a compression algorithm in accordance with a preferred embodiment of the present invention is depicted. The flowchart in **Figure 4** will be explained with reference to first data processing system 100 in **Figure 1**. The process begins at step 400, which depicts the initiation of the process. The process will typically be initiated by a signal from I/O module 128, which will, responsive to the preparations to transmit a data stream across

print controller 122 or network interface 120, set a flag in RAM unit 114 indicating that encoding system 130 needs to examine a datastream prior to transmission. The process then passes to step 402, which illustrates encoding system 130 ordering processor 112 to examine the datastream or a division of the datastream. The process next passes step 404, which depicts an attempt by encoding system 130 to discover the presence of an image data structure in the datastream. If no image data structure is present in the datastream, the process then passes to step 405, which depicts the datastream being transmitted by I/O module 128 across print controller 122 or network interface 120. The process then returns to step 402, which illustrates encoding system 130 ordering processor 112 to examine the datastream or a division of the datastream.

In step 404, if an image data structure is present in the datastream, the process then passes to step 406, which illustrates encoding system 130 extracting the image data structure from the datastream. The process next passes to step 408, which depicts encoding system 130 dividing the image into subregions. The process then passes to step 410, which illustrates encoding system determining if there are any additional subregions that require examination. If there are no additional subregions that require examination, the process next passes to step 412 which depicts encoding system 130 packaging the image for transmission. Packaging the image for transmission will typically involve the preparation of a packaged image data structure 300, which was described with reference to **Figure 3**. After packaged image data structure 300 is prepared, the process then passes to step 405, which depicts the datastream being transmitted by I/O module 128 across print controller 122 or network interface 120.

In step 410, if there are additional subregions that require examination, the process then passes to step 414, which illustrates encoding system 130 analyzing the next subregion of the image. On a first pass through step 408, the preferred embodiment will

typically conclude that there are additional subregions to analyze, because there will typically be at least a first subregion. This first subregion will be the "next" subregion analyzed in step 414. As a data structure, the subregion will typically be represented as a binary or hexadecimal series of digits. Analysis of the subregion will consist of a mathematical manipulation of the numerical representation of the subregion, such as the calculation of a checksum or other steps that can be used to increase the efficiency of comparison between two subregions. The process next passes to step 416, which depicts the encoding system comparing the subregion to a stored database of references in the stored data 115 to determine whether the subregion under analysis matches a previously stored reference. The references will be previously identified subregions that are associated with identifiers. The stored database of references can exist in a variety of forms, and multiple sources can contribute to the database of stored references. For instance, as each reference is identified in the analysis of an image according to the process of the preferred embodiment, the references and identifiers for that image can be stored in stored data 115 and can be used as part of the stored database of references. The references and identifiers can also be stored on fixed disk storage 116, and the references can be retained after the analysis of a first image data structure.

The references that are retained after the analysis of a first image data structure can then be retained as a symbol dictionary for use as part of the stored database of references in the analysis of all subsequent image data structures. The encoding system can also perform maintenance on the stored database of references, maintaining statistical counters to detail the frequency with which references in the symbol dictionary are employed and eliminating infrequently used references to save space or caching in RAM unit 114 any frequently used references from the symbol dictionary to save time in the encoding procedure. Frequently used references can also be preloaded on a receiving

machine and omitted from decoding table 304 of the packaged image data structure to further enhance transmission efficiency.

In step 416, if the subregion under analysis does not match any of the reference subregions, the process then passes to step 418, which illustrates the subregion under analysis being stored as a reference in stored data 115. An identifier is assigned to the reference at this time. The process next passes to step 420, which depicts the encoding system substituting the previously mentioned identifier for the numerical representation of the subregion in the image data structure. Multiple iterations of this process eventually produce reduced image 318. In step 416, if the subregion under analysis does match one of the reference subregions, the process then passes to step 420, which depicts the encoding system substituting the previously mentioned identifier for the numerical representation of the subregion in the image data structure. The process then passes to step 422, which illustrates the encoding system updating statistical counters to detail the frequency with which references in the symbol dictionary are employed and eliminating infrequently used references to save space or caching in RAM unit 114 any frequently used references from the symbol dictionary to save time in the encoding procedure. The process then returns to step 410.

With reference to **Figure 5**, a high-level flowchart for a process of decoding a packaged image data structure to reduce the bandwidth consumed in transmission of a datastream by employing a process of encoding image data within the datastream through a compression algorithm in accordance with a preferred embodiment of the present invention is illustrated.

The flowchart in **Figure 5** will be explained with reference to first display device 106 in **Figure 1**, though it could be explained as easily with reference to second data

processing system 102 or additional display device 110. The process begins at step 500, which depicts the initiation of the process. The process will typically be initiated by a signal from I/O module 144, which will, responsive to the receipt of a data stream across I/O controller 132, set a flag in RAM unit 134 indicating that decoding system 146 needs to examine a datastream that was received. The process then passes to step 502, which illustrates decoding system 146 ordering printing controller 136 to examine the datastream or a division of the datastream. The process next passes step 504, which depicts an attempt by decoding system 146 to discover the presence of a packaged image data structure in the datastream. If no packaged image data structure is present in the datastream, the process then passes to 505, which depicts the datastream being released by I/O module 144 to printing applications 142. The process then returns to step 502, which illustrates decoding system 146 ordering printing controller 136 to examine the datastream or a division of the datastream.

In step 504, if a packaged image data structure is present in the datastream, the process then passes to step 506, which illustrates decoding system 146 extracting the packaged image data structure from the datastream. The process next passes to step 508, which depicts decoding system 146 placing the references and identifiers from decoding table 304 into buffers in stored data 138 to create a stored database of references and identifiers. The references will be previously identified subregions that are associated with identifiers. The stored database of references can exist in a variety of forms, and multiple sources can contribute to the database of stored references. For instance, as each reference is identified in the previously described analysis of an image according to the process of the preferred embodiment, the references and identifiers for that image will typically be stored in decoding table 304 that is transmitted as a part of packaged image data structure 300. These references and identifiers will typically be used as part of the stored database of references. The references and identifiers can also be retained after

the analysis of a packaged image data structure and be stored for use in the analysis of later image data structures.

5 The references that are retained after the analysis of a first image data structure can be retained as a symbol dictionary for use as part of the stored database of references in the analysis of all subsequent image data structures. This will typically be accomplished by means of a signal in the packaged image data structure indicating that certain reference and identifier pairs are to be retained after the processing of the image in which they are used. The encoding system can also remotely perform maintenance on the stored database of references, maintaining statistical counters to detail the frequency with which references in the symbol dictionary are employed and sending signals to eliminate infrequently used references to save space. Frequently used references can also be preloaded on a receiving machine and omitted from decoding table 304 of the packaged image data structure to further enhance transmission efficiency.

10
15
20 The process then passes to step 510, which illustrates the decoding system determining if there are any additional identifiers present in reduced image 318. If there are additional identifiers, the process next passes to step 512, which depicts decoding system 146 replacing the identifiers in the reduced image with the references in stored data 138 to reproduce the original image. The process then returns to step 510, completing an iterative loop wherein the system replaces identifiers with references until all of the identifiers have been replaced.

25 In step 510, if there are no additional identifiers, the process then passes to step 514, which illustrates decoding system 146 decoding the image. This will involve preparing the image data structure in a format that the other software modules operating

on the display device can employ in other processes. The process then returns to step 505.

With reference to **Figure 6**, a module diagram representing message flow between the various software modules in a process of encoding image data within the datastream through a compression algorithm in accordance with the preferred embodiment of the present invention is depicted. The diagram depicts the flow of information between the encoding module, present on the sending system, and the decoding module, present on the receiving system. It also illustrates the intra-module traffic of messages between the various components of the encoding system and the decoding system, as they appear when the system is studied at the application level. The message flow in **Figure 6** will be explained with reference to encoding on first data processing system 100 and to decoding on second data processing system 102 in **Figure 1**, though it could be explained as easily with reference to first display device 106 or additional display device 110. The message flow begins at process initiation in step 400, when an input datastream 600 is received from I/O module 128 in operating system 124 of first data processing system 100. This input datastream flows into an examination buffer 602 within encoding system 601. I/O module 128 will simultaneously, responsive to the preparations to transmit a data stream across print controller 122 or network interface 120, set a flag in RAM unit 114 indicating that encoding system 130 needs to examine datastream 600 that is stored in examination buffer 602.

Next, after encoding system 130 orders processor 112 to examine the datastream or a division of the datastream in step 402, if encoding system 130 discovers the presence of an image data structure in the datastream in step 404, encoding system 130 will extract the image data structure from the datastream in step 406 and transfer image 604 from examination buffer 602 to image processor module 606. Encoding system 130 divides the image into subregions in step 408, analyzes the subregions in step 414, compares the

subregions to references in step 416, and, where it is discovered that a subregion does not match any previous references, an analyzed subregion is prepared as a reference 608, which is passed to a reference buffer 610 in step 418. Reference buffer 610 responds by assigning an identifier 612 to the reference and sending identifier 612 to image processor module 606, where the image processor module replaces the subregion with the identifier in step 420. The image processor module also maintains the subregion statistics in step 422. When all appropriate subregions have been replaced with references and step 410 indicates that there are no further subregions that require analysis, the image processor module sends reduced image 614 to packaging module 616 in step 412. Simultaneously, reference buffer 610 sends decoding table 618 to packaging module 616. Packaging module 616 then generates a header 302 and creates a packaged image data structure 300 by combining decoding table 618, header 302, and reduced image 614. Packaged image data structure 620 is then sent to a transmit buffer 622, where it is re-combined with datastream remainder 624, which was sent to transmit buffer 622 by examination buffer 602. Datastream remainder 624 comprises all information in the datastream other than image data structures that are the subject of the preferred embodiment.

The transmit datastream 626 is then sent by the transmit buffer to I/O module 128 in step 405. It travels across network interface 120 of first data processing system 100, across network 104, and across network interface 156 of second data processing system 102, to reach I/O module 162 in operating system 158 of second data processing system 102. I/O module 162 in operating system 158 of second data processing system 102 then passes transmit datastream 626 to receive buffer 628 in decoding system 630, initiating the decoding process in step 500.

The receive buffer 628 in decoding system 630 then examines the datastream or a division of it in step 502. In step 504, if a packaged image data structure is present in

the datastream, the receive buffer extracts the extracts the image data structure from the datastream in step 506. This consists of sending reduced image 632 to image processor module 634 and sending datastream remainder 636 to a storage buffer 638. Datastream remainder 636 comprises all information in the datastream other than image data structures that are the subject of the preferred embodiment. Decoding table 640 is then sent to reference buffer 642 and the references are stored in the reference buffer in step 508.

If, in step 510, image processor module 634 determines that there are identifiers present in reduced image 632, image processor module 634 replaces the identifiers in the reduced image with the references in stored data 138 to reproduce the original image in step 512. This is accomplished by image processor module 634 sending identifiers 644 from reduced image 632 to reference buffer 642 and reference buffer 642 replying to image processor module 634 by sending corresponding references 646 to image processor module 634. Image processor module 634 then replaces identifiers 644 with references 646 in step 512.

After the image data structure is completely reassembled by image processor module 634 in step 514, image data structure 648 is then transmitted to storage buffer 638, where it is recombined with datastream remainder 636 and released as an output datastream 650 to I/O module 162 in step 505.

Although aspects of the present invention have been described with respect to a computer system executing software that directs the functions of the present invention, it should be understood that present invention may alternatively be implemented as a program product for use with a data processing system. Programs defining the functions of the present invention can be delivered to a data processing system via a variety of signal-bearing media, which include, without limitation, non-rewritable storage media

(e.g., CD-ROM), rewritable storage media (e.g., a floppy diskette or hard disk drive), and communication media, such as digital and analog networks. It should be understood, therefore, that such signal-bearing media, when carrying or encoding computer readable instructions that direct the functions of the present invention, represent alternative embodiments of the present invention.